ON THE ASYMPTOTIC EFFICIENCY OF A SEQUENTIAL PROCEDURE FOR ESTIMATING THE MEAN[1]

By Norman Starr[2]

Columbia University and University of Minnesota

1. Introduction. Let the independent, identically distributed random variables

$$X_1, X_2, \cdots \tag{1}$$

be $N(\mu, \sigma^2)$ with μ unknown and $0 < \sigma < \infty$. Define for $n \geqq 2$,

$$\bar{X}_n = n^{-1} \sum_{i=1}^n X_i, \qquad S_n^2 = (n-1)^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2, \tag{2}$$

and suppose that for fixed $s, t > 0$ the loss incurred in estimating μ by $\bar{X}_n$ from a sample of fixed size n is

$$L_n = A|\bar{X}_n - \mu|^s + n^t \qquad (A > 0), \tag{3}$$

with risk

$$\nu_n(\sigma) = E_\sigma L_n = AE_\sigma|\bar{X}_n - \mu|^s + n^t. \tag{4}$$

When σ is *known* the problem of finding the value of n, say n^0, for which the risk (4) is a minimum is perfectly straightforward; let $\nu(\sigma)$,

$$\nu(\sigma) = \nu_{n^0}(\sigma) = \min_{n>0} \nu_n(\sigma), \tag{5}$$

denote the minimum risk. On the other hand, *in ignorance of* σ no procedure based on a fixed number n of observations of (1) will minimize (4) simultaneously for all $0 < \sigma < \infty$. Accordingly, the possibility of utilizing a sample of random size N determined by a certain sequential rule $\mathfrak{R}$ to be specified later, will be considered. In analogy with (3) the loss using $\mathfrak{R}$ is for fixed $s, t > 0$ and N,

$$L_N = A|\bar{X}_N - \mu|^s + N^t \qquad (A > 0), \tag{6}$$

with risk

$$\bar{\nu}(\sigma) = E_\sigma L_N = AE_\sigma|\bar{X}_N - \mu|^s + E_\sigma N^t. \tag{7}$$

It would seem to be of considerable practical importance to compare the values of $\nu(\sigma)$ and $\bar{\nu}(\sigma)$ for values $0 < \sigma < \infty$ of the parameter upon which these functions depend. For, either it will turn out that ν and $\bar{\nu}$ do not differ appreciably for any value of σ, in which case a very useful and easily applied statistical procedure will have been justified, or in the contrary case a horrible example of the dangers of "optional stopping" will have been exposed.

Received 19 November 1965.

[1] Research supported in part by National Science Foundation Grant NSF-GP-3694 at Columbia University and NSF-GP-3813 at the University of Minnesota.

[2] Now at Carnegie Institute of Technology.

The impetus for considering $\mathfrak{R}$ was provided by H. Robbins [4], who computed $\bar{\nu}$ for a number of values of σ in the special case $s = t = 1$ of the loss criterion (3), verifying that the difference between $\bar{\nu}$ and ν is negligible at the values of σ considered. Presumably $\mathfrak{R}$, for other, if not all, specializations of the loss (3), will exhibit this desirable property for moderate values of σ.

The purpose of this article is to establish a condition on $\mathfrak{R}$ (specificially, the condition on the *starting sample size* m of the procedure) for which the *risk efficiency*,

$$\eta(\sigma) = \bar{\nu}(\sigma)/\nu(\sigma), \tag{8}$$

i.e. the ratio of the risk using $\mathfrak{R}$ in ignorance of σ, to the minimum risk using a sample of fixed size n^0 when σ is known, converges to 1 as σ becomes infinite. This condition turns out to be simply

$$m > s^2/(s + 2t) + 1. \tag{9}$$

Thus *for example* in the special case $s = 2$, $t = 1$ of the loss criterion (3), i.e. squared-error loss and unit cost of sampling, $\mathfrak{R}$ is asymptotically risk efficient if and only if it requires that $N \geqq 3$.

2. Preliminaries. Suppose σ is *known*. We wish to find the value n^0 (a function of σ) for which the risk (4) is a minimum. Observe that (3) may be rewritten as

$$L_n = (A\sigma^s/n^{s/2})|n^{\frac{1}{2}}(\bar{X}_n - \mu)/\sigma|^s + n^t, \qquad 0 < \sigma < \infty. \tag{10}$$

Since for fixed n,

$$Z = n^{\frac{1}{2}} \cdot (\bar{X}_n - \mu)/\sigma \tag{11}$$

is $N(0, 1)$, the risk (4) is simply

$$\begin{aligned} \nu_n(\sigma) &= (A\sigma^s/n^{s/2})(2/\pi)^{\frac{1}{2}} \textstyle\int_0^\infty z^s e^{-z^2/2}\, dz + n^t \\ &= (2/s)K\sigma^s/n^{s/2} + n^t, \end{aligned} \tag{12}$$

where

$$K = K(A, s) = (s/2)A\Gamma((s + 1)/2)2^{s/2}/\Gamma(\tfrac{1}{2}) \tag{13}$$

depends on A and s only. Treating $n \geqq 0$ as a continuous variable, we have

$$\partial\nu_n(\sigma)/\partial n = -K\sigma^s/n^{s/2+1} + tn^{t-1}, \tag{14}$$

which, when set equal to zero, has the unique positive root

$$n^0 = (K\sigma^s/t)^{2/(s+2t)}; \tag{15}$$

since from (12)

$$\lim_{n\to 0} \nu_n(\sigma) = \lim_{n\to\infty} \nu_n(\sigma) = \infty, \tag{16}$$

n^0 is the value of $n \geqq 0$ which minimizes (4). Accordingly, we have by simple algebra from (15) and (12) that the minimum risk (5) is

$$\nu(\sigma) = ((2/s)t + 1)(n^0)^t. \tag{17}$$

Suppose now that σ is *unknown*. We will consider the following sequential procedure:

Let K be defined by (13), and let $\{k_n, n \geqq 1\}$ be any sequence of positive constants such that

$$\lim_{n\to\infty} k_n = K; \tag{18}$$

then in analogy with (15) we define the sequential rule $\mathfrak{R}$: Observe the sequence (1) term by term, stopping with X_N, where

$$N \text{ is the first integer } \quad n \geqq m \quad \text{ such that } \quad n \geqq (k_{n-1}S_n^s/t)^{2/(s+2t)}, \tag{19}$$

with the *starting sample size* $m \geqq 2$ a fixed integer.

As in *for example* [4] and [5], define the independent $N(0, 1)$ random variables

$$W_n = (\textstyle\sum_{i=1}^n X_i - nX_{n+1})/\sigma[n(n+1)]^{\frac{1}{2}} \qquad (n \geqq 1), \tag{20}$$

and

$$V_k = \textstyle\sum_{j=1}^{k-1} W_j^2 \qquad (k \geqq 2). \tag{21}$$

Observing that

$$V_n = (n-1)S_n^2/\sigma^2 \qquad (n \geqq 2), \tag{22}$$

it is easily seen that (19) may be rewritten in the form

$$N \text{ is the first integer } \quad n \geqq m \quad \text{ such that } \quad V_n \leqq l(n, \sigma), \tag{23}$$

where for $n \geqq 2$, $0 < \sigma < \infty$,

$$l(n, \sigma) = (t/k_{n-1})^{2/s}(n-1)n^{(s+2t)/s}/\sigma^2; \tag{24}$$

then the probability distribution of N is defined for $n \geqq m$ by

$$p_n(\sigma) = P(N = n) = P(V_{k+1} \leqq l(k+1, \sigma) \quad \text{for} \quad k = n-1, \text{ but not for any } \quad m-1 \leqq k \leqq n-1). \tag{25}$$

We propose to evaluate the performance of $\mathfrak{R}$ in terms of the risk $\bar{\nu}$ defined by (7). As in [4] and [5], observe that for fixed $n \geqq 2$, $\bar{X}_n$ is independent of the vector $(W_1, \cdots, W_{n-1})$; hence L_n and the event $\{N = n\}$ are independent for all $n \geqq m$, and

$$\begin{aligned} \bar{\nu}(\sigma) &= \textstyle\sum_{n=m}^\infty p_n(\sigma)E_\sigma(L_N \mid N = n) \\ &= \textstyle\sum_{n=m}^\infty p_n(\sigma)\nu_n(\sigma) \\ &= \textstyle\sum_{n=m}^\infty p_n(\sigma)((2/s)K\sigma^s/n^{s/2} + n^t) \\ &= (2/s)K\sigma^s E(N^{-s/2}) + EN^t, \qquad 0 < \sigma < \infty, \end{aligned} \tag{26}$$

and

(27) $$\eta(\sigma) = \bar{\nu}(\sigma)/\nu(\sigma)$$
$$= (2/s)K\sigma^s E(N^{-s/2})/((2/s)t + 1)(n^0)^t + EN^t/((2/s)t + 1)(n^0)^t,$$

simplifying, with some algebra, to

(28) $$\eta(\sigma) = ((2/s)t + 1)^{-1}((2/s)t(n^0)^{s/2}E(N^{-s/2}) + E(N/n^0)^t).$$

To show that $\mathfrak{R}$ is asymptotically risk efficient, that is, that

(29) $$\lim_{\sigma\to\infty} \eta(\sigma) = 1,$$

we could verify that for arbitrary fixed ω,

(30) $$EN^\omega \sim (n^0)^\omega \quad \text{as} \quad \sigma \to \infty,$$

for then (29) follows immediately from (28). However, it is apparent that the validity of (29) depends in some way on the size m of the starting sample of $\mathfrak{R}$. For suppose that the terms of the sequence (1) have a very large variance. Then if s is "large" relative to t in a specialization of the loss criterion (3), the portion of loss assignable to error of estimate will effectively "dominate" that associated with the cost of sampling. This suggests that the risk may blow up "out of all proportion" if the possibility that sampling is terminated very early is permitted. Accordingly, we establish in the next section a necessary and sufficient condition on the starting sample size m for which (29) holds; this condition amounts to the verification of (30) for $\omega = t$ and, when (9) is satisfied, for $\omega = -s/2$.

3. Results.

THEOREM 1. *With n^0 defined by* (15) *and with N defined by* (19) (*or* (23))

(31) $$\operatorname{plim}_{\sigma\to\infty} (N/n^0) = 1.$$

Expression (31) may be proved directly by extending the methods of [3], or may be treated as a corollary to Lemma 1 of [2].

THEOREM 2. *For $\omega > 0$ fixed*

(32) $$\lim_{\sigma\to\infty} E(N/n^0)^\omega = 1.$$

PROOF. Let $0 < \epsilon < 1$ be given, and define

(33) $$\alpha = (1 - \epsilon)^{\omega^{-1}} n^0;$$

then

(34) $$EN^\omega = \sum_{n=m}^{\infty} n^\omega P(N = n) \geqq \alpha^\omega P(N \geqq \alpha).$$

Therefore,

(35) $$E(N/n^0)^\omega \geqq (1 - \epsilon)P(N \geqq \alpha),$$

and from Theorem 1

(36) $$\liminf_{\sigma\to\infty} E(N/n^0)^\omega \geqq 1 - \epsilon.$$

Also, with $0 < \epsilon < 1$ given, define

(37) $$\beta = (1 + \epsilon)^{\omega^{-1}} n^0;$$

then

(38) $$EN^\omega = \sum_{n=m}^{\infty} n^\omega P(N = n) \leqq (\beta + 1)^\omega P(N \leqq \beta + 1) + T(\beta),$$

where

(39) $$T(\beta) = \sum_{n \geqq \beta+1} n^\omega P(N = n).$$

Rewriting (38) in the form

(40) $$E(N/n^0)^\omega \leqq ((\beta + 1)/n^0)^\omega P(N \leqq \beta + 1) + T(\beta)/(n^0)^\omega,$$

it is clear that for σ sufficiently large, *if*

(41) $$T(\beta) < \mathcal{K}$$

where $\mathcal{K}$ is a constant independent of β, *then* Theorem 1 together with (40) imply

(42) $$\limsup_{\sigma \to \infty} E(N/n^0)^\omega \leqq 1 + \epsilon$$

which, with (36), proves (32). Hence, the proof is complete if we verify (41).

To prove (41), note from (23) that

(43) $$\{N = n\} \subset \{V_{n-1} > l(n - 1, \sigma)\};$$

then,

(44) $$T(\beta) = \sum_{n \geqq \beta} (n + 1)^\omega P(N = n + 1) \leqq \sum_{n \geqq \beta} (n + 1)^\omega P(V_n > l(n, \sigma)).$$

Define for $n \geqq 2$, $0 < \sigma < \infty$,

(45) $$h(n, \sigma) = l(n, \sigma)/(n - 1).$$

Inasmuch as the sequence $\{k_n\}$ satisfies (18), we can always choose $\sigma(\epsilon)$ so large that for $\sigma > \sigma(\epsilon)$,

(46) $$k_{n-1} \leqq (1 + \epsilon)^{(s+2t)/2\omega - s\tau/2} \cdot K, \qquad 0 < \tau < (s + 2t)/s\omega, \quad n \geqq \beta.$$

Then it is easily seen that for $\sigma > \sigma(\epsilon)$,

(47) $$h(n, \sigma) \geqq (1 + \epsilon)^\tau, \qquad \tau > 0, \quad n \geqq \beta;$$

hence, from (44) and (47), for $\sigma > \sigma(\epsilon)$,

(48) $$T(\beta) \leqq \sum_{n \geqq \beta} (n + 1)^\omega P(V_n/(n - 1) > h(n, \sigma)) \leqq \sum_{n \geqq \beta} (n + 1)^\omega P(V_n/(n - 1) - 1 > (1 + \epsilon)^\tau - 1), \quad \tau > 0,$$

which, with the $2m$th moment Markov inequality ($m \geqq 1$), gives

(49) $$T(\beta) \leqq \sum_{n \geqq \beta} (n + 1)^\omega \cdot E(V_n - (n - 1))^{2m}/((1 + \epsilon)^\tau - 1)^{2m}(n - 1)^{2m}, \quad m \geqq 1, \quad \tau > 0.$$

Observe that $V_k (k \geqq 2)$, originally defined by (21), is a chi-squared random variable with $k - 1$ degrees of freedom. It can be proved by an induction on p that for $p = 1, 2, \cdots$,

$$(50) \qquad E(V_{k+1})^{r+p} = (k + 2r) \cdots (k + 2r + 2p - 2)E(V_{k+1})^r$$

where $k \geqq 1$, $r \geqq 1$ are integers. Then it is easy to establish the fact that $E(V_{k+1} - k)^{2m}$ = a polynomial "$P^m(k)$" in k of order m ($m \geqq 1, k \geqq 1$), by expanding the argument, performing an induction on m, and applying (50). Now let

$$(51) \qquad m = \{\omega\} + 1,$$

where $\{x\}$ = smallest inteter $> x$; then from the previous fact, (49), and (51), for $\sigma > \sigma(\epsilon)$,

$$(52) \qquad \begin{aligned} T(\beta) &\leqq \textstyle\sum_{n \geqq \beta} [(n + 1)^\omega P^{\{\omega\}+1}(n)/(n - 1)^{2\{\omega\}+2}] \\ &\leqq C \textstyle\sum_{n \geqq \beta} n^{-(1+\gamma)} \leqq C^1 \beta^{-\gamma} \leqq \mathcal{K} \qquad (0 < \gamma \leqq 1), \end{aligned}$$

verifying (41), and completing the proof of (32). □

We have already commented that the condition (9) on the starting sample size m of $\mathfrak{R}$ for which (30), and therefore (29), holds, is reflected in the limiting behavior of the negative moments of the sampling variable N. To formalize, we define for $s, t > 0$ (as always fixed),

$$(53) \qquad \begin{aligned} a(m, \omega) &= 2^{-(m-3)/2} m^{-\omega}/(m - 1)\Gamma((m - 1)/2), \\ b(\omega) &= (K/t)^{2\omega/(s+2t)}, \\ l(n) &= \sigma^2 l(n, \sigma), \qquad n \geqq 2, \end{aligned}$$

and let

$$(54) \qquad \mathfrak{a}(m, \omega) = a(m, \omega) b(\omega) l(m)^{(m-1)/2},$$

where the fixed integer $m \geqq 2$ denotes as usual the starting sample size of $\mathfrak{R}$; then we have

THEOREM 3. *For $\omega > 0$ fixed*

$$(55) \qquad \begin{aligned} \lim_{\sigma\to\infty} (n^0)^\omega EN^{-\omega} &= 1, && \textit{for} \quad m > 2s\omega/(s + 2t) + 1, \\ &= 1 + \mathfrak{a}(m, \omega), && \textit{for} \quad m = 2s\omega/(s + 2t) + 1, \\ &= \infty, && \textit{for} \quad m < 2s\omega/(s + 2t) + 1. \end{aligned}$$

PROOF. From (15) and (24) we have

$$(56) \qquad (n^0)^\omega l(m, \sigma)^{(m-1)/2} = b(\omega) l(m)^{(m-1)/2} \sigma^{2s\omega/(s+2t)-(m-1)};$$

thus, the theorem follows immediately *provided*

$$(57) \qquad \lim_{\sigma\to\infty} (n^0)^\omega EN^{-\omega} = 1 + a(m, \omega) \cdot \lim_{\sigma\to\infty} \{(n^0)^\omega l(m, \sigma)^{(m-1)/2}\}.$$

To prove (57), let $0 < \epsilon < 1$ be given and with α and β defined by (33) and (37), respectively, define for $\omega > 0$ fixed

$$\begin{aligned}\pi_1 &= m^{-\omega}P(N = m),\\ \pi_2 &= \beta^{-\omega}P(m < N \leqq \beta),\\ \pi_3 &= \textstyle\sum_{m+1 \leqq n \leqq \alpha} n^{-\omega}P(N = n),\\ \pi_4 &= \alpha^{-\omega}P(N \geqq \alpha).\end{aligned} \tag{58}$$

(A) Observe that

$$EN^{-\omega} \geqq \pi_1 + \pi_2 . \tag{59}$$

Recalling that V_k $(k \geqq 2)$ is a chi-squared random variable with $k - 1$ degrees of freedom, we have from (25)

$$\begin{aligned}\pi_1 &= m^{-\omega}P(V_m \leqq l(m, \sigma))\\ &= m^{-\omega}/\Gamma((m-1)/2)2^{(m-1)/2} \textstyle\int_0^{l(m,\sigma)} x^{(m-1)/2-1}e^{-x/2}\,dx\\ &\geqq a(m, \omega)e^{-l(m,\sigma)/2}l(m, \sigma)^{(m-1)/2}.\end{aligned} \tag{60}$$

Noting that for fixed $n \geqq 2$,

$$\lim_{\sigma\to\infty} l(n, \sigma) = 0, \tag{61}$$

we obtain from (59) and (60)

$$\begin{aligned}\liminf_{\sigma\to\infty} (n^0)^{\omega}EN^{-\omega} \geqq{}& a(m, \omega) \lim_{\sigma\to\infty} \{(n^0)^{\omega}l(m, \sigma)^{(m-1)/2}\}\\ &+ (1 + \epsilon)^{-1} \lim_{\sigma\to\infty} P(m < N \leqq \beta)\end{aligned} \tag{62}$$

which, with Theorem 1, gives

$$\begin{aligned}&\liminf_{\sigma\to\infty} (n^0)^{\omega}EN^{-\omega}\\ &\quad\geqq a(m, \omega)\cdot\lim_{\sigma\to\infty} \{(n^0)^{\omega}l(m, \sigma)^{(m-1)/2}\} + 1 - \delta \quad (0 < \delta = \delta(\epsilon) < 1).\end{aligned} \tag{63}$$

(B) Also, observe that

$$EN^{-\omega} \leqq \pi_1 + \pi_3 + \pi_4 . \tag{64}$$

We have first of all

$$\begin{aligned}\pi_1 &= m^{-\omega}P(V_m \leqq l(m, \sigma))\\ &= m^{-\omega}/\Gamma((m-1)/2)2^{(m-1)/2} \textstyle\int_0^{l(m,\sigma)} x^{(m-1)/2-1}e^{-x/2}\,dx\\ &\leqq a(m, \omega)l(m, \sigma)^{(m-1)/2}.\end{aligned} \tag{65}$$

Next, remarking from (23) that

$$\{N = n\} \subset \{V_n \leqq l(n, \sigma)\}, \tag{66}$$

we obtain

$$\begin{aligned}\pi_3 &\leqq \textstyle\sum_{m+1\leqq n\leqq\alpha} n^{-\omega}P(V_n \leqq l(n,\sigma))\\
(67)\qquad &= \textstyle\sum_{m+1\leqq n\leqq\alpha} [n^{-\omega}/\Gamma((n-1)/2)2^{(n-1)/2}] \int_0^{l(n,\sigma)} x^{(n-1)/2-1}e^{-x/2}\,dx\\
&= \textstyle\sum_{m+1\leqq n\leqq\alpha} [n^{-\omega}/\Gamma((n-1)/2)2^{(n-1)/2}\\
&\qquad\cdot \int_0^{l(n,\sigma)} (x^{(n+\rho)/2-1}e^{-x/2})x^{-(\rho+1)/2}\,dx, \quad 0<\rho<1.\end{aligned}$$

With $h(n,\sigma)$ defined by (45), observe that since the sequence $\{k_n\}$ satisfies (18) we can always choose $\sigma(\epsilon)$ so large that for $\sigma > \sigma(\epsilon)$,

$$(68)\qquad h(n,\sigma) \leqq 1-\xi \quad \text{for all} \quad n \leqq \alpha,$$

where $0 < \xi = \xi(\epsilon) < 1$. Observe further that there exists a value ρ, $0 < \rho < 1$, such that for $n \geqq 3$, $\sigma > \sigma(\epsilon)$,

$$(69)\qquad n+\rho-2 > (n-1)(1-\xi) \geqq (n-1)h(n,\sigma) = l(n,\sigma).$$

Then from (67), since $X^{(n+\rho)/2-1}e^{-x/2}$ achieves its maximum at $n+\rho-2$, we have

$$\begin{aligned}\pi_3 &\leqq \textstyle\sum_{m+1\leqq n<\alpha} [n^{-\omega}/\Gamma((n-1)/2)2^{(n-1)/2}]\cdot l(n,\sigma)^{(n+\rho)/2-1}e^{-l(n,\sigma)/2}\\
&\qquad\cdot\int_0^{l(n,\sigma)} x^{-(\rho+1)/2}\,dx\\
(70)\qquad &= [2/(1-\rho)] \textstyle\sum_{m+1\leqq n\leqq\alpha} [n^{-\omega}/\Gamma((n-1)/2)2^{(n-1)/2}]\\
&\qquad\cdot l(n,\sigma)^{(n-1)/2}e^{-l(n,\sigma)/2}\\
&= [2/(1-\rho)] \textstyle\sum_{m+1\leqq n\leqq\alpha} [n^{-\omega}(n-1)^{(n-1)/2}/\Gamma((n-1)/2)2^{(n-1)/2}\\
&\qquad\cdot [h(n,\sigma)e^{-h(n,\sigma)}]^{(n-1)/2}, \qquad \sigma > \sigma(\epsilon).\end{aligned}$$

Moreover,

$$(71)\qquad \Gamma((n-1)/2) \geqq \int_0^{(n-1)/2} x^{(n-1)/2-1}e^{-x}\,dx \geqq ((n-1)/2)^{(n-1)/2-1}e^{-(n-1)/2}.$$

Let

$$(72)\qquad \Delta(n,\sigma) = h(n,\sigma)e^{(1-h(n,\sigma))};$$

then it is easily verified from (68) that for $\sigma > \sigma(\epsilon)$,

$$(73)\qquad \Delta(n,\sigma) \leqq (1-\xi)e^{\xi} < 1, \qquad n \leqq \alpha, \quad 0 < \xi < 1.$$

Hence, from (70), (71) and (72), we obtain

$$\begin{aligned}\pi_3 &\leqq [2/(1-\rho)] \textstyle\sum_{m+1\leqq n\leqq\alpha} n^{-\omega}[(n-1)/2]\Delta(n,\sigma)^{(n-1)/2}\\
(74)\qquad &= 2/(1-\rho) \textstyle\sum_{m+1\leqq n\leqq\alpha} n^{-\omega}[(n-1)/2]\Delta(n,\sigma)^{m/2}\Delta(n,\sigma)^{(n-m-1)/2}\\
&\leqq \sigma^{-m}\cdot 2e(1-\rho)^{-1} \textstyle\sum_{m+1\leqq n\leqq\alpha} n^{-\omega}[(n-1)/2]\\
&\qquad\cdot l(n)^{m/2}\Delta(n,\sigma)^{(n-m-1)/2}, \quad \sigma > \sigma(\epsilon),\end{aligned}$$

which, with (73) and the *ratio rule* for series convergence, gives for $\sigma > \sigma(\epsilon)$

$$\pi_3 \leqq \mathcal{G}/\sigma^m, \tag{75}$$

where $\mathcal{G}$ is a constant independent of σ. Therefore, from (64), (65), and (75), we obtain for $\sigma > \sigma(\epsilon)$,

$$(n^0)^\omega EN^{-\omega} \leqq a(m, \omega)(n^0)^\omega l(m, \sigma)^{(m-1)/2} + (\mathcal{G}/\sigma^m)(n^0)^\omega + (1 - \epsilon)^{-1}P(N \geqq \alpha). \tag{76}$$

From definitions (15) and (24), we have

$$\lim_{\sigma\to\infty} (n^0)^\omega l(m, \sigma)^{(m-1)/2} < \infty \Rightarrow \lim_{\sigma\to\infty} [\mathcal{G}(n^0)^\omega/\sigma^m] = 0; \tag{77}$$

hence, from (76), (77), and Theorem 1, we obtain finally

$$\limsup_{\sigma\to\infty} (n^0)^\omega EN^{-\omega} \leqq a(m, \omega) \lim_{\sigma\to\infty} \{(n^0)^\omega l(m, \sigma)^{(m-1)/2}\} + 1 + \delta' \quad (0 < \delta' = \delta'(\epsilon) < 1), \tag{78}$$

which, with (63), completes the proof of (57), and *a fortiori*, (55). □

Let $\eta(\sigma)$, $0 < \sigma < \infty$, be defined by (8), and let $\mathfrak{a}(m, \omega)$ be defined by (54), where the fixed integer $m \geqq 2$ denotes the *starting sample size* of $\mathfrak{R}$; then we have the main result:

COROLLARY. *For* $s, t > 0$ *fixed*,

$$\lim_{\sigma\to\infty} \eta(\sigma) \tag{79}$$
$$\begin{aligned} &= 1, && \textit{for} \quad m > s^2/(s + 2t) + 1, \\ &= 1 + (2/s)t\mathfrak{a}(m, s/2)/((2/s)t + 1), && \textit{for} \quad m = s^2/(s + 2t) + 1, \\ &= \infty, && \textit{for} \quad m < s^2/(s + 2t) + 1. \end{aligned}$$

The corollary follows from Theorem 2 (with $\omega = t$), Theorem 3 (with $\omega = s/2$), and (28). □

4. Remarks. Suppose we modify the cost of sampling in (3) so that for fixed $s > 0$ the loss is

$$L_n^* = A\,|\bar{X}_n - \mu|^s + \log n \qquad (A > 0); \tag{80}$$

then, by methods analogous to those of Section 2 it can be shown that when σ is *known*

$$\nu_n^*(\sigma) = (2/s)K\sigma^s/n^{s/2} + \log n \qquad \text{(risk)}, \tag{81}$$

$$(n^0)^* = K^{2/s}\sigma^2 \qquad \text{(minimizing value of } n \geqq 0), \tag{82}$$

and

$$\nu^*(\sigma) = 2/s + \log (n^0)^* \qquad \text{(minimum risk)}. \tag{83}$$

When σ is *unknown* we consider the sequential procedure $\mathfrak{R}^*$, which (with $\{k_n\}$ defined to be any sequence of positive constants satisfying (18)) is, in analogy with (82),

(84) $\mathfrak{R}^*$: Observe the sequence (1) term by term, stopping with X_{N^*}, where N^* is the first integer $n \geqq m$ for which $n \geqq k_{n-1}^{2/s} S_n^{\ 2}$,

with the starting sample size $m \geqq 2$ a fixed integer.

Using $\mathfrak{R}^*$ in ignorance of σ it can be confirmed that

$$\nu^*(\sigma) = (2/s)K\sigma^s E((N^*)^{-s/2}) + E(\log N^*) \qquad \text{(risk)} \tag{85}$$

and from (83) and (85), for $0 < \sigma < \infty$,

$$\eta^*(\sigma) = (2/s)((n^0)^*)^{s/2}E((N^*)^{-s/2})/(2/s + \log (n^0)^*) + E(\log N^*)/(2/s + \log (n^0)^*) \quad \text{(risk efficiency of } \mathfrak{R}^*). \tag{86}$$

With $a(m, \omega)$, $m \geqq 2$, defined in (53), let

$$\mathfrak{a}^*(m, \omega) = a(m, \omega)[(K/k_{m-1})^{2/s}(m - 1)m^{s/2}]^\omega; \tag{87}$$

then by the methods of Section (3), we can prove:

THEOREM 1*. *With* $(n^0)^*$ *defined by* (82) *and with* N^* *defined by* (84)

$$\text{plim}_{\sigma\to\infty} (N^*/(n^0)^*) = 1. \tag{88}$$

THEOREM 2*. *For* $\omega > 0$ *fixed*

$$\lim_{\sigma\to\infty} E(N^*/(n^0)^*)^\omega = 1. \tag{89}$$

THEOREM 3*. *For* $\omega > 0$ *fixed*

$$\begin{aligned} \lim_{\sigma\to\infty} ((n^0)^*)^\omega EN^{*-\omega} &= 1, && \textit{for} \quad m > 2\omega + 1, \\ &= 1 + \mathfrak{a}^*(m, \omega), && \textit{for} \quad m = 2\omega + 1, \\ &= \infty, && \textit{for} \quad m < 2\omega + 1. \end{aligned} \tag{90}$$

Thus, from (86), (89), and (90), we have

COROLLARY*. *For* $s > 0$ *fixed*

$$\begin{aligned} \lim_{\sigma\to\infty} \eta_\sigma^* &= 1, && \textit{for} \quad m \geqq s + 1, \\ &= \infty, && \textit{for} \quad m < s + 1. \end{aligned} \tag{91}$$

These results, in addition to their interest for the problem of obtaining a point estimate of μ (with minimum risk), can be related in a simple way to the problem of estimating μ by an interval of fixed width (with prescribed confidence) when, as in the former problem, the experimenter is observing terms of the sequence (1).

It is easy to show, see *for example* [3], that when σ is *known*, an interval estimate

$\{\bar{X}_n - d, \bar{X}_n + d\}$ for μ of width $2d$ $(d > 0)$ and with coverage probability at least α $(0 < \alpha < 1)$ is guaranteed, provided the sample size n is chosen so that

$$n \geqq a^2\sigma^2/d^2 \qquad (0 < \sigma < \infty), \tag{92}$$

where a is the solution of

$$(2\pi)^{-\frac{1}{2}} \int_{-a}^{a} e^{-t^2/2}\, dt = \alpha. \tag{93}$$

When σ is *unknown*, the following sequential procedure has been studied [1], [2], [3], [5]:

Let $\{a_n, n \geqq 1\}$ be any sequence of positive constants for which

$$\lim_{n\to\infty} a_n = a, \tag{94}$$

and in analogy with (92) define the sequential rule

(95) $\mathfrak{R}_*$: Observe the sequence (1) term by term, stopping with X_{N_*}, where N_* is the first integer $n \geqq m_*$ such that $n \geqq a_{n-1}^2 S_n^2/d^2$, with $m_* \geqq 2$ a fixed integer.

Then form the interval

$$I_{N_*} = (X_{N_*} - d, X_{N_*} + d) \tag{96}$$

which has the required width.

The limiting behavior of N_* and of its first moment have received considerable attention, for example [1], [2], [3]; it has been proved that

$$\text{plim}_{\sigma\to\infty} (d^2 N_*/a^2\sigma^2) = 1, \tag{97}$$

from which it follows easily that

$$\lim_{\sigma\to\infty} P(\mu \in I_{N_*}) = \alpha \qquad (\text{asymptotic consistency of } \mathfrak{R}_*), \tag{98}$$

and moreover,

$$\lim_{\sigma\to\infty} E\, d^2N_*/a^2\sigma^2 = 1 \qquad (\text{asymptotic efficiency of } \mathfrak{R}_*). \tag{99}$$

(In fact, (97), (98), and (99) hold irrespective of the underlying common distribution of the terms of (1).)

Observe now that the results of this section relating to the problem of point estimation enable us to describe the limiting behavior of *all* of the moments of N_*, the sampling variable for the procedure $\mathfrak{R}_*$. For, suppose we take as our loss criterion (80), and let

$$k_n = (a_n/d)^s, \qquad n = 1, 2, \cdots; \tag{100}$$

that for preassigned $0 < \alpha < 1$ and $d > 0$ this is permissible follows from the fact that the sequence $\{k_n\}$ may be chosen in any way such that (18) holds, where K (which, for fixed $s > 0$, depends only on (arbitrary) A) is arbitrary. Therefore,

$$K = \lim_{n\to\infty} k_n = \lim_{n\to\infty} (a_n/d)^s = (a/d)^s. \quad (101)$$

Hence, rewriting (82) in the form

$$(n^0)^* = \{(a/d)^s\}^{2/s}\sigma^2 = a^2\sigma^2/d^2, \quad (102)$$

and (84) in the form

(103) N^* is the first integer $n \geqq m$ for which

$$n \geqq \{(a_{n-1}/d)^{2/s}\}^s S_n^2 = a_{n-1}^2 S_n^2/d^2,$$

we see by comparing (92) and (95) with (102) and (103), respectively, that the performance of $\mathfrak{R}_*$ and $\mathfrak{R}^*$ (for the special choice (100) of the $\{k_n\}$) is identical. Therefore, for $0 < \alpha < 1$, $d > 0$ preassigned, we have

THEOREM 1_*. *With a defined as the solution of* (93)

$$\operatorname{plim}_{\sigma\to\infty} (d^2N_*/a^2\sigma^2) = 1; \quad (104)$$

reproving (97).

THEOREM 2_*. *For $\omega > 0$ fixed*

$$\lim_{\sigma\to\infty} E(d^2N_*/a^2\sigma^2)^\omega = 1. \quad (105)$$

THEOREM 3_*. *For $\omega > 0$ fixed, and $\mathfrak{A}^*(m, \omega)$ defined by* (87),

$$\begin{aligned} \lim_{\sigma\to\infty} E(d^2N_*/a^2\sigma^2)^{-\omega} &= 1, && \textit{for} \quad m_* > 2\omega + 1 \\ &= 1 + \mathfrak{A}^*(m_*, \omega), && \textit{for} \quad m_* = 2\omega + 1 \\ &= \infty, && \textit{for} \quad m_* < 2\omega + 1. \end{aligned} \quad (106)$$

In particular, note that (99) follows as a corollary to Theorem 2_*. Moreover, we observe from Theorem 2_* that the asymptotic sampling variability of $\mathfrak{R}_*$ may be partially characterized by

$$\begin{aligned} \lim_{\sigma\to\infty} \operatorname{var} (d^2N_*/a^2\sigma^2) &= \lim_{\sigma\to\infty} E(d^2N_*/a^2\sigma^2)^2 \\ &\quad - \lim_{\sigma\to\infty} (E(d^2N_*/a^2\sigma^2))^2 = 1 - 1 = 0. \end{aligned} \quad (107)$$

Finally, we remark that all the results of this article hold if the sequential procedures we have considered are modified in such a way that the experimenter is permitted to terminate sampling *only* with *odd*-numbered observations (in which case the starting sample size is a fixed odd integer $\geqq 3$). The performance of such modified rules has been evaluated for moderate values of σ by Robbins [4] with relation to the problem of point estimation, and by the author [5] with relation to the problem of fixed-width interval estimation.

Acknowledgment. I am greatly indebted to Professor Herbert Robbins, who suggested this problem and under whose guidance the research was conducted, for his invaluable advice and continuing encouragement.

I also wish to thank Professor Leon J. Gleser for his many helpful comments and suggestions.

REFERENCES

[1] ANSCOMBE, F. J. (1953). Sequential estimation. *J. Roy. Statist. Soc. Ser. B* **15** 1–21.

[2] CHOW, Y. S. and ROBBINS, H. (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean. *Ann. Math. Statist.* **36** 457–462.

[3] GLESER, L. J., ROBBINS, H., and STARR, N. (1964). Some asymptotic properties of fixed-width sequential confidence intervals. Columbia University Report.

[4] ROBBINS, H. (1959). Sequential estimation of the mean of a normal population. *Probability and Statistics* (Harald Cramér Volume), Almqvist and Wiksell, Uppsala. 235–245.

[5] STARR, N. (1966). The performance of a sequential procedure for the fixed-width interval estimation of the mean. *Ann. Math. Statist.* **37** 36–50.